Readability of Online Hearing Related Information in Russian

Alina Filatova

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Audiology

University of Canterbury,

Department of Communication Disorders

2018

Table of Contents

Acknowledgment

I would firstly like to thank the Department of Communication Disorders for giving me this incredible opportunity to study audiology. It has been challenging but extremely worth it. Secondly I would like to thank the wonderful staff both academic and administrative for looking after us and supporting us throughout this journey. It would not have been possible to get through it without such a great team. To Jonny, Nicole, Virginia and Bethany, without your constant clinical advice and support I would not be where I am now and I am very grateful to have had the pleasure of working with each one of you. A massive thank you also to Rebecca, you have been the most incredible, supportive and helpful thesis supervisor I could have ever asked for. I apologise for my lack of research skills and hope I didn't cause you too much agony in this long writing process. I definitely would not have made it without you.

Lastly, to my incredible classmates, we did it. Through the late nights, the tears, the laughter, and the many blocks of chocolate. I couldn't have asked for a better group of people to go on this journey with, so, well done everyone and good luck for the future as audiologists.

Abstract

Purpose: To describe the readability of hearing-related Internet-information for the Russian language and to compare these results with the results of the Cloze test administered to native speakers of Russian.

Method: The readability of Russian online hearing-related information was examined by using the Google country-coded top-level domains (ccTLDs) in six countries where Russian is spoken by at least 5% of the population. These countries included; Russia, Kazakhstan, Ukraine, Belarus, Moldova, and Turkmenistan. The key search terms used were; "deterioration of hearing"- "Ухудшение слуха", "hearing impairment"- "Нарушения слуха", and "hearing loss"- "потеря слуха", and "снижение слуха". The top 10 webpages that adhered to the inclusion and exclusion criteria of this study were retrieved for each search term in each ccTLD. Duplicates were removed and a total of 39 webpages were used for analysis. Readability was analysed using four readability formulas: Flesch-Kincaid (F-K), Simple Measure Of Gobbledygook (SMOG), Dale-Chall readability formula, and Coleman-Liau Index. The Cloze test was also used and was administered to 10 participants. It was performed to make a comparison between the webpages that had the highest and lowest readability scores.

Results: The analysis of the 39 webpages in this study concluded that the overall readability levels were high, meaning the general public needed to have at least 13-16 years of education to understand and comprehend the information available on the webpages. The lowest readability level was found when using the Dale-Chall formula, which gave a reading grade level (RGL) of 7.41 which showed adequate readability. The highest readability level was found when using the F-K formula which gave a RGL of 24.66 which showed poor readability. The mean RGL of all formulas ranged from 13.02-16.44 concluding that all

analysed webpages had significantly higher RGLs than the recommended RGL of six. The Cloze test produced results comparable to the high RGLs found from the formulas. The webpage with the lowest reading grade level gave a mean Cloze score of 64.08% compared to the webpage with the highest reading grade level which gave a mean Cloze score of 45.78% showing that the lower reading grade level webpage was easier to complete.

Conclusion: Readability of the Russian language has not been well documented, and currently there have not been any studies looking at the readability of hearing-related information on the Internet of different countries in the Russian language. As many people now turn to the Internet for information, it is important that this information can be easily understood by the public. Clinicians and healthcare professionals should also be able to recommend appropriate websites and give their patients high quality materials that are easy enough for them to understand. This is important to achieve greater client satisfaction and improved decision making for the public.

List of abbreviations

ANOVA= Univariate analysis of variance

ASHA= American Speech-Language-Hearing Association

ccTLD= country-coded Top Level Domain

CRIE= Chinese Readability Index Explorer

F-K= Flesch-Kincaid

GBD= Global burden of Disease

HHIE= Hearing Handicap Inventory for the Elderly

HON= Health On the Net

HRQoL= Health-related quality of life

IP = Internet Protocol

ITU= International Telecommunications Union

PEW= Pew Research Center

RA= Readability analyser

RGL= Reading grade level

SMOG= Simple Measures of Gobbledygook

URL= Uniform resource locator

USA= United States of America

US= United States

WHO= World Health Organisation

List of Tables

1 Introduction

1.1 Study overview

The Internet has become an important resource for both health consumers and health care professionals (Porter & Edirippulige, 2007). Today's public is able to access the Internet on many platforms such as computers, mobile phones, laptops and even gaming devices (Statistics NZ- Tatauranga Aotearoa, 2012). Gartner, Inc., a leading research and advisory company based in United States of America (USA), forecasted that by the end of 2017, 8.4 billion "connected things- laptops, mobiles etc." will be used worldwide (Egham, 2017). This estimate will be up 31% from 2016 and has the potential to reach 20.4 billion by 2020 (Egham, 2017).

Having easy access to the Internet on different devices has also made it convenient to access health related information by the public at any time. A United States (US) national survey was conducted by Baker and colleagues in 2001 which surveyed 4764 individuals aged 21 years and over who reported to be Internet users. The survey found that approximately 40% of respondents used the Internet to look for health related advice or information (Baker et al., 2003). Another US based survey conducted by Hesse and colleagues in 2005 looked at how common it was for the public to use the Internet to search for health-related information as well as their level of trust in health information sources. This survey had a total of 6369 individuals aged 18 years and over. The survey concluded that 63.7% of the Internet users in this sample looked for health related information either for themselves or others (Hesse et al., 2005). Both people with health conditions and people who need information quickly are able to access thousands of healthcare materials from their own homes. According to Statistics New Zealand- Tatauranga Aotearoa (2012), 2.8 million New

Zealanders used the Internet. They also reported that 50% of people who went on the Internet looked for information on health services.

Obtaining the correct health information on the Internet quickly, conveniently and in private helps people make informed decisions on their next move (Porter & Edirippulige, 2007). However, it has been found that the quality of health-related information on the Internet is variable and that health care professionals should take the time to direct patients to sources that have good quality information about what they are looking for (Shepperd, Charnock, & Gann, 1999). The Porter and Edirippulige study in 2007 examined patterns of how parents of deaf children sought out hearing-related information on the Internet in Australia. The study targeted parents of children with a permanent hearing loss aged from birth to 21 years. A total of 166 individuals completed the survey. The study found that 87% of these individuals tended to use generic search engines (Google) for finding information, followed by parents visiting websites that they felt specialised in hearing loss, and websites that were recommended by other parents of deaf children.

An individual's level of education can have an impact on the level of Internet use (Porter & Edirippulige, 2007). Individuals with low literacy skills have been found to have worse health outcomes than individuals who are literate as they tend to use health services less, and are unable to make informed decisions from the online health information they find (Berkman et al., 2011; McInnes & Haglund, 2011). Unfortunately, websites on the Internet tend to use many complex jargon words, which are difficult for low literate individuals to understand and read (McInnes & Haglund, 2011).

There has been extensive research conducted on the readability of health-related information in which results showed above average RGLs of the assessed health information (Atcherson et al., 2014; Wang et al, 2013; Berland et al., 2001; Walsh & Volsko, 2008;

Cheng & Dunn, 2015). Readability of hearing-related information has also been researched and multiple studies conducted. For example Kelly and Kahn in 1991 and Kelly in 1996 analysed hearing related information and too found RGLs that were equivalent to college reading levels.

High RGL and readability scores of health and hearing information is not only a problem in the English language. Traditional Chinese has also been found to have high readability levels of hearing-related information on the Internet (Hsu, 2017)**.** The readability of languages such as Spanish, English and Chinese has been researched to see if the high readability trends are evident. However, many other languages have not yet been assessed.

The readability of health-related information in Russian has not been widely studied to date. Since Russian is one of the top ten spoken languages in the world, it is vital that the readability of Russian language websites are analysed. This study will delve into the online hearing information available to the public on the Internet to see if the information is also at an RGL that is too high for consumers to understand.

1.2 Overview of hearing impairment

A hearing impairment refers to a total or partial loss of a person's ability to hear causing the individual to not be able to hear as well as someone with normal hearing (WHO, 2017). There are many causes of hearing impairment in both adults and children, which can occur in different parts of the auditory system. The human auditory system is made up of the outer ear, middle ear, inner ear, auditory nervous pathways and higher cortical areas. Each part of the auditory system and their structures work differently in the way they perceive sound and make sense of it to get the most information for the listener.

The outer ear, specifically the auricle (pinna), is used for the localisation of sounds (Katz, 2015). The middle ear is located in the temporal bone and consists of the tympanic membrane (ear drum) and the three ossicles (malleus, incus and stapes) which transmit the mechanical energy of the sound signals from the ear drum through to the inner ear by vibrating the ossicular chain (Katz, 2015). The inner ear is fluid filled and consists of the auditory and balance organs. The stapes footplate of the middle ear connects to the oval window which when moved produces pressure waves in the cochlea, these pressure waves transform the mechanical energy into neural impulses which are sent to the brain via the auditory nerve (Dillon, 2012). The auditory system is essentially a chain reaction where when one structure is stimulated, the rest follow. When a hearing impairment comes into play, this chain can be impacted causing problems with either the sounds entering the ear or the inability of the sounds being processed. Hearing impairments can arise through different situations. They can be congenital: present before or at birth, genetic: run in the family, or can be acquired later in life.

1.2.1 Types of hearing impairments

There are three types of hearing losses that both adults and children can have: conductive, sensorineural or mixed. A conductive hearing loss involves an abnormality of the outer or middle ear and is usually caused due to a loss of transmission of sound to the inner ear (Yueh et al., 2003; Strawbridge et al., 2000). The sound is unable to travel through the structures or is attenuated and therefore does not make it to the inner ear at the level that it should. Individuals with conductive losses may still have normal underlying hearing. Conductive hearing losses can occur due to many different reasons such as cerumen build up,

otitis media, otosclerosis, tympanosclerosis or microtia, sometimes these types of losses are relatively easy to correct and often cause a temporary hearing loss (Strawbridge et al., 2000).

A sensorineural hearing loss is when the inner ear (cochlea and associated structures) is damaged, preventing sounds from getting to the brain. Usually high frequency sounds are affected first as the basiliar membrane is tonotopically organised starting from high frequencies and going too low (Weinstein, 1994). This makes the hearing loss affect many speech sounds that are used for clarity causing a person extreme difficulty when there is background noise. Such a hearing loss is very common with older people and is called presbycusis where the hair cells naturally deteriorate; other causes of a sensorineural hearing loss could be noise exposure, and certain medications such as aminoglycoside antibiotics (Strawbridge et al., 2000). Sensorineural hearing losses are permanent as the hair cells do not undergo regeneration after being damaged (Shepherd & Hardie, 2001).

A mixed hearing loss occurs when there is both a problem with the outer or middle ear and the inner ear. This type of loss can be due to a conductive hearing loss such as otosclerosis but left untreated and the inner ear is then affected or it could be two separate problems.

As mentioned previously, hearing loss can either be congenital or acquired. Congenital hearing losses can be due to complications at birth such as lack of oxygen or an infection the mother may have contracted during pregnancy such as rubella or any ototoxic drugs that had entered the bloodstream (Fisch, 1973). Genetic hearing losses are also common. An example of this has been documented in four families worldwide who have mitochondrial inherited deafness due to a mutation in the A7445G gene (Love & Bird, 2013).

This genetic mutation leads to a hearing loss that can develop between the ages of 5-6 years of age starting as a mild sensorineural hearing loss and progressing to a severe or profound loss between the ages of 20-30 (Love & Bird, 2013). Acquired hearing loss can manifest from several different reasons and can occur at any age. Some of these causes can be due to: noise exposure, through work or hobbies without wearing hearing protection, ageing, head traumas, and viral infections among other things (ASHA, n.d., WHO, 2017).

1.2.2 Interventions for hearing impairment

Different intervention strategies are available for the different types of hearing losses. Sensorineural hearing losses usually can be helped with the use of a hearing aid, depending on the degree of hearing loss. Cochlear implants are also an option for people who do not benefit from a hearing aid. Correcting conductive losses can include micro-suctioning if the problem is cerumen build up, or corrective surgery if the damage is more severe such as otosclerosis (Briggs & Luxford, 1994). Depending on the degree of hearing loss in any case, a conventional hearing aid is typically enough to help the person hear well. Hearing impairment impacts many people all over the world and is a global issue. Hearing impairment can be a disabling condition for many individuals.

1.2.3 Prevalence of hearing impairment

The global prevalence of hearing loss was first estimated in 1985, with 42 million (0.9%) of the world's population thought to be affected (Smith, 2003). Within 15 years this number had significantly increased, affecting 120 million people in 1995 and then further increased to 250 million (4.2%) people by 2001 (Smith, 1998; Mathers Smith & Concha, 2003). By 2012 the number of people affected by a disabling hearing loss had further increased to 360 million (roughly over 5%) of the world population, of which 32 million are children (WHO, 2012). World Health Organisation (WHO) refers to a disabling hearing loss

as a permanent, unaided hearing threshold greater than 40dB in the better hearing ear in adults, and greater than 30dB in children (WHO, 2012). Hearing loss has a considerable impact on a person's day to day communication ability therefore it is a communication disorder as it can disrupt the use of spoken or written language.

1.2.3.1 Prevalence of hearing impairment in the younger population

Hearing loss has a huge impact on communication and can affect an individual's ability to function day to day. Hearing loss can be apparent in both adults and children and is highly prevalent. Hearing loss has been identified as the most common birth defect in children as estimates by the National Institute on Deafness and Other Communication Disorders have shown that around 16,000 to 18,000 babies and toddlers are identified with hearing loss per year in the US (Madell & Flexer, 2014). A study conducted by Niskar and colleagues in 1998 looked at the prevalence of hearing loss among children in the US aged between 6 to 19 years. The sample size of this study was 6,166 children. They completed the audiometry in the examination centre of the Third National Health and Nutrition Examination Survey conducted between 1988 and 1994. This study concluded that 14.9% of US children had low or high frequency hearing loss of at least 16 dB HL in one or both ears (Niskar et al, 1998). Another US based study conducted by Shargorodsky and colleagues, (2010) looked at the prevalence of hearing loss of adolescents (aged 12 to 19 years) in the US. Their study compared results from 1988 to 1994 study with results from 2005-2006. They concluded that the number of adolescents with hearing loss has increased in 2005-2006. This increase went from 14.9% in the years 1988 to 1994, up to 19.5% in 2005 to 2006.

1.2.3.2 Prevalence of hearing impairment in the adult population

The Global Burden of Disease (GBD) Project study had estimated that hearing impairment in adults was the third leading cause of disability (Stevens et al., 2013). A study

conducted by Feder and colleagues, (2015) found that in the years of 2012-2013 there were 4.6 million (19%) Canadians aged between 20 and 79 years who had a hearing loss. The researchers of this study reported that before audiometric data was completed, the participants were asked to do a self-report of their hearing. Less than 4% of these people reported to experience hearing difficulties, when in fact 19% had at least a mild hearing loss. It has been shown that childhood hearing impairment can reduce the ability to communicate with others later in life as well as causing poor language acquisition which can lead to the inability of interpreting speech sounds, have educational disadvantages and can lead to social isolation (Stevens et al., 2011; Olusanya, Neumann, & Saunders, 2014). In 2008, Agrawal and colleagues showed that the prevalence of hearing loss could be growing due to an ageing population and the increase of noise exposure.

1.2.3.3 Prevalence of hearing impairment in Russia

Information about the prevalence of hearing impairment in the Russian population is limited with few data available. By land mass, Russia is the largest country in the world with a population of over 144 million people recorded at the end of March 2016 (Federal State Statistics Service, 2016). The Russian language is the 8th most commonly spoken language in the world with around 280 million speakers worldwide (Simmons & Charles, 2017).

1.2.4 Impact of hearing impairment

Hearing impairment first and foremost affects the individual themselves. It has been shown that people with mild hearing losses can struggle with verbal language processing which can severely affect a person's life (Agrawal et al., 2008). An individual's hearing loss can also impact family members, friends and co-workers as it can alter the person's ability to communicate with others effectively therefore affecting those relationships (Slawinski, Hartel, & Kline, 1993). The impact of hearing impairment in children can be devastating.

Developmental delay of speech and language are one of the main concerns of childhood hearing loss as it can cause learning difficulties as well as problems through schooling and friendships (Tyler, Tye-Murray, & Gantz, 1991). These problems can then manifest later in life and lead to social isolation, depression and even cause issues in their physical functioning (ASHA, 2016).

1.2.4.1 Impact of hearing impairment on adults

Adults with a hearing loss also experience negative effects in different aspects of their lives especially if the onset of the hearing loss occurs later in life. Depression, poor self-esteem, loneliness and diminished functioning are all negative consequences that have been found to be associated with hearing loss in the elderly (Dalton et al., 2003; Monzani et al., 2008; Strawbridge et al., 2000; Chen, 1994). Although their speech and language may have developed normally, adults with a hearing loss will still tend to ask for a repeat of what was said or focus a lot on lip-reading (Summerfield, 1992). This can be very strenuous and some people may find it easier to not communicate at all. A study conducted by Chen in 1994 looked at the impact of hearing loss in adults over the age of 65 years using four self-reporting questionnaires that were completed by the individuals. The study found that hearing handicap – referring to the individual's response to the hearing loss – had a significant correlation with both loneliness and low self-esteem. These negative consequences not only impact the person with the hearing loss, but can also impact the people around them.

1.2.4.2 Impact of hearing impairment on significant others

The people affected the most would be the closest to that person such as: a spouse, parents, close family, friends, caregivers and colleagues (Kamil & Lin, 2015). They are the ones who can be deeply affected as they communicate with the person who has a hearing loss on a daily basis. A systematic review was carried out in 2015 by Kamil and Lin to explore the

literature on how a hearing loss of an individual affects their communication partners. 24 studies were examined. Many of them concluded that there was a decrease in social functioning of the communication partners due to the hearing loss of the person who was hard of hearing, and that the communication difficulties were a source of stress which led to lower relationship satisfaction for the communication partners. The communication partners are required to put more effort in when speaking to the person with the hearing loss, they tend to speak slower, and articulate their words clearly, they need to think about these and other factors when communicating with the person who has a hearing loss. For some people this can take a toll on them and become too difficult, therefore abandoning all communication.

Parents whose children have a hearing loss also experience some negative consequences such as anxiety, stress and sadness. When parents find out that their child has a hearing loss, some experience shock, an emotional crisis and lose their confidence in their ability to know how to act towards the child (Mason & Mason, 2007). Despite the initial shock and stress, parents are able to turn to health care professionals so early intervention can take place for the progression of communication skills of that child (Mason & Mason, 2007). It has been shown that family members are the most common reason why a person with a hearing loss looks for help (Mahoney et al., 1996). These family members or significant others have also been used to help gauge the persons hearing disability before and after any treatment option used (Newman & Weinstein, 1988).

1.2.5 Management of hearing impairment

When a hearing impairment is untreated, it can cause many negative issues as described above. However, there are ways that a person with a hearing loss can reduce and sometimes completely avoid the negative consequences. Hearing rehabilitation most commonly involves amplification by fitting hearing aids, using hearing assisted devices and

communication strategies (Yueh et al., 2003). Each option is dependent on the individual as well as on the degree of hearing loss. The benefit of using amplification is most commonly measured by using self-reports (Chisolm et al., 2007). A systematic review conducted by Chisolm and colleagues (2007) looked at the health-related quality of life (HRQoL) benefits of amplification in adults with a sensorineural hearing loss. The review included 16 studies that used HRQoL assessments which examine the degree at which people rate their health status to affect their self-perception of daily functioning and well-being. The Hearing Handicap Inventory for the Elderly (HHIE) developed by Ventry and Weinstein in 1982, was found to be the most commonly used self-reported measure. It is a 25-item questionnaire that gives an assessment of the persons hearing loss as thought of by them. The systematic review found that when using the HHIE as an outcome measure, individuals reported reductions in the emotional and social impacts of hearing loss after the use of hearing aids (Chisolm et al., 2007).

Kochkin and Rogin (2000) assessed the benefits on quality of life of hearing aid use by conducting a survey of 2069 individuals with hearing loss and their family members. A questionnaire based on previous quality of life questionnaires was developed by the researchers and given to both the individual with the hearing loss and their family members. The researchers found that the use of hearing aids improved these individual's social, emotional, psychological and physical well-being. This in turn improved social relationships, cognitive functioning and reduced frustration and communication difficulties. The use of hearing aids or any other assistive devices can have a positive impact on both the individual with the hearing loss and their communication partners and overcome other negative consequences that can manifest from having a hearing loss.

1.3 Health information

Consumers and their families are becoming more interested in finding health information themselves to maintain good health as well as understand risk factors and learn preventative measures (Brashers, Goldstein, & Hsieh, 2002). Health information is available to the public and can be found through many different modes. Traditionally healthcare providers such as doctors were the main source of health information; however, friends, family members, and media including magazines, television, newspapers, and radio have also been used as sources for obtaining health information (Brashers, Goldsmith, & Hsteh, 2002; Pennbridge, Moya, & Rodrigues, 1999; Cotten & Gupta, 2004). Although today's public is better informed and has access to health information through many different and convenient means, healthcare providers are still a trusted and reliable source of health information as they are seen as experts, and can sometimes have the most influence on a person's decisions (Pennbridge, Moya, & Rodrigues, 1999; Hesse et al., 2005). Individuals who have an informational orientation are interested in and motivated to actively seek out health information (Dutta-Bergman, 2004). These people tend to be the ones who come to the healthcare providers prepared with the information they found. Having the different means of accessing health information can help the consumer make informed decisions about their own health especially when a health provider cannot immediately be contacted.

1.3.1 Modes of health information

Health information can be given to the consumer in many different forms. Verbal health information can be obtained orally from a healthcare provider, from health videos on the Internet or from call centres specifically designed for health-related issues. Written health information can be obtained from books, brochures, pamphlets and information sheets that

can be found in clinics and hospitals, or given to the individual by the healthcare provider. Retrieving health information from the Internet has become a popular source for consumers and their families (McInnes & Haglund, 2011). According to Fox and Jones (2009), almost half of all the people who search for information online find that it influences their health decisions. Obtaining the correct health information on the Internet quickly, conveniently and in private helps people make informed decisions on their next steps (Porter & Edirippulige, 2007). Both people with health conditions and people who need information quickly are able to access thousands of healthcare materials from the convenience of their own homes.

1.3.2 Online health information

Obtaining health information off the Internet provides advantages such as anonymity among people who do not want their condition to be known until they know a bit about it themselves, access to large amounts of information, the ability to look for specific information and also the ability to interact with others who may have the same health problems/questions (Cline & Haynes, 2001; Cotten & Gupta, 2004). Using the Internet for health information can also have disadvantages. This can include hindering the patient-healthcare provider relationship. It has been noted that some physicians already experience patients coming to them requesting certain procedures, medications or tests because they found the information online (Heese et al., 2005). This can in turn cause distrust on the patient's behalf if the physician does something differently to what they had read (Heese et al., 2005). Other disadvantages can include the general disorganization of websites, the use of technical language and searching difficulties (Cotten & Gupta, 2004).

1.3.2.1 Global access to online health information

The Pew Research Center (PEW; 2017) have documented that in the early 2000s about half of all adults in America were already using the Internet. By the end of 2016 nine-

in-ten American adults were using the Internet. Despite the above listed disadvantages, the entire world still uses the Internet whether it is for entertainment purposes or looking for health information. According to the Reuters insight business report of 2003, health information was searched by 53% of Americans on the Internet. McMullan (2006) reported that there had been a significant increase in searching for online health information among Americans in the last decade. It was estimated that in 2002, 110 million Americans (80%) had searched for health information (McMullan, 2006). A survey completed by the International Telecommunications Union (ITU) estimated that at the end of 2016, there would be just fewer than 3.5 billion individuals world-wide using the Internet. In 2017 according to the Internet World Stats, there are 3.7 billion people worldwide using the Internet. The majority of Internet users are from Asia with half of the population recorded to be using the Internet. The second region to have the most internet users globally is Europe with 17% (Internet World Stats, 2017).

1.3.2.2 Access to online health information in Russian speaking populations

As of 2017 the Russian Federation had a population of roughly 144 million (World Bank Group, 2017). Of this population as of March 2017, 104.5 million were Internet users. Internet penetration is defined as the percent of the population that is using the Internet. In Russia 72.9% of the population was using the Internet in March 2017 (Internet World Stats, 2017). However, Russia is not the only Russian speaking country therefore the use of the Internet for Russian speakers could be even higher. Russian is spoken in the countries of the former Soviet Union, these include: Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Tajikistan, Turkmenistan, Uzbekistan and Ukraine (Learning Russian, n.d). According to PEW (2015), 63% of Internet users in Russia searched for health information. The most popular use of the Internet was for socializing (PEW, 2015). Although the Internet is an easy tool to use, the information found

can sometimes be inaccurate or hard to comprehend therefore leading the individual to confusion or inaccurate decisions (Cotton & Gupta, 2004).

1.4 Quality of online health information

Health information although useful can sometimes be misleading and could cause people to make wrong decisions. It has been found that the quality of health-related information on the Internet is variable and that health care professionals should take the time to direct patients to sources that have good quality information about what they are looking for (Shepperd, Charnock, & Gann, 1999). Health information on the Internet can sometimes be inaccurate as there is no oversight of what information websites can publish. People who are illiterate or have poor medical knowledge can struggle with both understanding the information they come across as well as struggle to judge whether this information is reliable or not (Atcherson et al., 2014). Although there are no set rules on what can be published on the Internet, there are ways of determining what the RGL of a text or website is through readability formulas.

1.4.1 Quality control of online health information

There has been an increase in the number of health resources that have become easily accessible due to the Internet. Previously, there was a cost for information production resulting in fewer resources available. Now however, almost anyone can post on the Internet as there is no prerequisite of what content can and cannot be published on the Internet (Metzger, 2007). Most websites do not have any oversight of the information that is being posted and some do not have any research to back up their facts. This makes it extremely difficult for the public to know what information is credible and trustworthy and what is not. Previous research into the quality of online health information, has found that the quality varies between websites and some information is inaccurate (Morhan-Martin, 2004). Some

organizations have put forward recommendations as to what can be published however not everyone follows these recommendations

1.4.2 Certification of compliance

The Health on the Net (HON) foundation is a non-profit, non-governmental organisation. Its focus is on promoting and guiding healthcare professionals to deploy useful and reliable health information online (HON, 2017). The HONcode is a code of conduct that is applied for health/medical websites to help create a standard of reliability of the medical and health information that is available on the Internet to patients, professionals, and the general public (HON, 2017). A website that is granted the HONcode must abide by eight key principles. These principals help in making sure the website is up to standard.

While the Internet can be an easy tool to use for finding information, it has been shown that many websites are hard for people to read and understand and therefore problems can arise when trying to make informed decisions. In 2002, the European commission created a code of conduct which had a quality criteria set in place for health websites; these criteria included things such as; honesty, privacy, transparency, and authority among other things (McInnes & Haglund, 2011). This was done so that it was easier for Internet users to find authentic and correct health related information. HON is a non-profit, non-governmental organisation that promotes and guides online health information to be useful, reliable and to protect people from misleading health information (HON, 2016), A website that has a HONcode is one that has been certified by the organisation and tells the audience that the website provides quality, objective and transparent medical information (HON, 2016). However, even with the HONcode, websites on the Internet can still be hard to read and understand.

1.5 Importance and use of health information

Demand for consumer health information is very high for both public and professional consumption (Deering & Harris, 1996). People use health information to maintain good health as well as find information to learn and understand about risks and preventative measures (Brashers, Goldsmith, & Hseih, 2002). However, the use of health information found by the public can vary in quality and comprehensibility.

1.5.1 Health literacy

Health literacy refers to a person's ability to find, interpret and use the information and health services to effectively make decisions about their health and well-being (Ministry of Health NZ; Manatu Hauora, n.d). It has been shown that functional literacy – the ability to have basic reading and writing skills – is important as it is these skills that make people literate allowing them to have a higher level of control over their everyday life events as they are more informed (Nutbeam, 2008). The notion of health literacy now requires more than just basic reading and writing skills. To truly be considered health literate other factors have been included such as a person's ability to access, understand fully and be able to use the health information from a variety of sources (Batterham et al., 2016).

Having a high level of health literacy allows the person to make informed decisions about their health. It has been shown that people with low health literacy have worse health outcomes as they tend to use health services less than people with higher health literacy (Berkman et al., 2011). According to Sultz (2006), there were 98 million Americans that used the Internet to search for health information in 2006. As discussed previously the Internet has become one of the most popular sources for obtaining health care information however it does not mean it is the easiest source to understand, people still need to have a high level of health literacy.

Literacy skills have been measured in terms of a person's RGL, which is defined by Doak, Doak, and Root, (1996) as the average reading skill that is achieved each year at school in the American public-school system. Reports have shown that the average RGL for the adult population in America is at the 8th to 9th grade level, however, one out of five Americans can only read at a 5th grade level or lower (Doak, Doak, & Root, 1996). This means that they cannot understand most newspapers, or health-related information. If a text is written for independent or unassisted use, then the RGL is usually higher than if the text is used for learning purposes, such as in a classroom (Dubay, 2004). A person who has a RGL below 5 is thought to have low literacy and someone with a reading grade level at the 5th grade and above is thought to be literate (Doak, Doak, & Root, 1996).

Older adults are most at risk of having literacy difficulties (Friedman, Hoffman-Goetz, & Arocha, 2009). This population is also the most likely to be seeking out the health information as they tend to have more health issues, or are more likely to look after a child or others with health issues (Deering & Harris, 1996). Therefore, it is important the health information available is at a level that is comprehendible for this population.

1.5.2 Readability

Websites on the Internet tend to use many complex and jargon words, which is why as mentioned above, it can be hard for people to understand and read these websites. Readability is defined as how easy it is to read and understand written information (Laplante- Lévesque et al., 2012). Jargon and lengthy sentences can reduce the readability of a particular text, meaning it is harder for someone to understand. Dubay (2004, p.3), referred to readability as "what makes some texts easier to read than others." Health-related information needs to be at a RGL that is easy enough to be understood by most people so that they are able to make informed decisions and not misinterpret information. Readability of information related to

hearing loss has previously been analysed and has been found to be at high RGLs. A study conducted by Kelly and Kahn (1991), looked at the readability of clinic forms. They requested clinic forms that clients receive from universities, hospitals and private practices. They found that over half of the documents sampled from speech and hearing facilities needed a high school level of education or higher to be able to understand the documents. The readability of hearing aid brochures has also been looked into in a study by Kelly (1996), which also found that 58% of all documents analysed were at a college level. This again shows that many people who are reading these brochures may not fully understand what they mean and therefore will not read further.

A systematic review conducted by Laplante-Lévesque & Thorén, (2015), looked at the readability of Internet information on hearing that can be accessed by people with a hearing impairment or their significant others. This review examined eight studies and found that, on average, people with a hearing impairment, or their significant other, needed 9 to 14 years of education in order to read and understand the Internet information on hearing (Laplante-Lévesque & Thorén, 2015). A study that was included in the systematic review conducted by Atcherson et al. (2014) looked at evaluating the readability of audiology and speech and language pathology public content on the American Speech-Language-Hearing Association (ASHA) website. This study found that the ASHA website articles that are accessible by the public, 225 exceeded the recommended RGL and on average were found to be at a 9th RGL (Atcherson et al., 2014).

This is not only a problem in the English language. Traditional Chinese has also been found to have high RGLs of hearing related information on the Internet (Hsu, 2017). Hsu 2017 looked at the readability of traditional Chinese through the Jing and CRIE 1.0 readability formulas as there were no other readability formulas available. He concluded that with the CRIE 1.0 readability formula, 25% of websites had a RGL greater than 6, and the

Jing formula showed that 81% of websites had a RGL greater than 6. The readability of health-related information in Russian has not been widely studied. This is the reason why this study will be analysing the readability of Russian websites.

1.6 Readability analysis

The use of readability analyses has become an easy and popular way of measuring readability of health care materials as well as other things such as safety manuals and even legal materials such as working contracts. Readability formulas are used to help the writer of a particular text by finding possible grammatical and stylistic issues to help maintain a consistent level of the document (McCallum & Peterson, 1982). There are two main ways of completing readability analyses and these include computational analyses or non-computational analyses.

1.6.1 Computational analysis

Statistical techniques are used to develop readability formulas through computational analysis. Originally readability formulas were calculated by hand as there was limited use of computers. It has been documented that as early as 1963 attempts were made to start computerizing these formulas (McCallum & Peterson, 1982). It has been established that the common variables that are used in readability formulas are semantic and syntactic measures (Dubay, 2004). Semantic measures refer to meaning, such as difficulty with vocabulary, and syntactic measures refer to sentence structure, such as average sentence length. These have been found to best predict text difficulty (Dubay, 2004). Computational readability analyses use algorithms that are derived from computational linguistics and used in conjunction with prediction models that create for a more robust and accurate analysis of text difficulty (Collins-Thompson, 2014). Using computational analyses has the advantage of being a fast and easy way of calculating readability scores. As well as this, computational analyses also

evolve as vocabulary evolves and become better at analysing the more complex semantic resources that are becoming available online (Collins-Thompson, 2014).

1.6.1.1 *Textual features in English language readability formulas*

Textual features are used to indicate the reading difficulty of a text. Word and sentence features are the most commonly used features in readability formulas. Alphabetic languages tend to have similar textual features and therefore readability formulas in other languages such as Spanish follow similar principals as English formulas (Spaulding, 1956). The first words we learn are usually the shortest and easiest this also makes them the most frequently used words. Johnson (1946) as cited in Dubay (2004) found that 24 life stories that were written by university students, of the 67,200 words used, 25% consisted of ten words as follows: the, I, and, to, was, my, in, of, a and it. It has been shown that half of all written materials are made up of the first 100 most frequent words (Fry et al., 1993). If people do not have an extensive vocabulary it is harder for them to understand certain texts that may not contain words that are used most frequently.

A strong predictor of someone's reading skills as well as verbal development has been shown to be the knowledge of words (Dubay, 2004). The length of words as well as syllable counts has also been used as a predictor of text difficulty. Shorter words, with fewer syllables are easier and quicker to read than words that are longer and therefore have more syllables. Sentences have also been used to predict text difficulty. Just as with shorter words, shorter sentences also make a text easier to read and understand. Long sentences tend to have complex grammatical structures and can be a strain on the reader as they have to remember several parts of the sentence before being able to make it into a meaningful whole (McLaughlin, 1969).

1.6.1.2 Textual features in Russian language readability formulas

Teachers of both German and Russian used word counts to match texts with students as they too found that frequently used words are more familiar and easier to use (Thorndike, 1921). To date, there has been no specific readability formula developed for Russian, instead some of the already widely used readability formulas have been adapted to accommodate for the Russian language.

1.6.2 Non-computational analysis

Computational analyses are useful in that they are performed on computers and not by hand, therefore making it easier and faster. Non-computational analyses however, have been shown to be better at taking into account more difficult linguistic features (Klare, 1954). This makes them a useful tool to use with computational analysis to get the most accurate and reliable outcome. An example of a non-computational analysis is a comprehension test and is described below.

1.6.2.1 Comprehension tests

A comprehension test can show someone's reading ability as well as how much they understand what they are reading. Reading texts involves both being able to decode the text and comprehend what is written (Keenan, Betjemann, & Olsen, 2008). One of the main comprehension tests that aid in measuring readability of a text is the Cloze test which offers a convenient and valid way of measuring both the reading comprehension of people or the comprehension difficulties of texts (Bormuth, 1969). The Cloze test is described below.

1.6.2.2 Cloze test

The Cloze test was developed by Wilson Taylor in 1953 as a tool for measuring readability. Readers fill in each blank with the word they think belongs in the blank, using

pragmatic, syntactic and semantic clues available in the remaining text (Gemoets et al., 2004). The procedure is founded on the theory that readers at the appropriate level of the text should be able to provide closure to the passage as they construct meaning by correctly filling in the blanks (Stephens, 2000). In this procedure, participants are given a passage from a text that have words missing at periodic intervals (e.g. every fifth word) and are asked to fill in the missing words (Gemoets et al., 2004).

As well as assessing readability of texts, The Cloze test also assesses a person's comprehension ability by determining how much the person understands from reading a specific passage (Alderson, 1979; Miller et al., 2009). The passages are required to be between 250-300 words, the first and last sentences should remain intact and the blank lines should be similar in length to the deleted word. The responses are scored as either correct or incorrect. The person's ability to correctly fill in the missing words is how comprehension is measured as opposed to measuring the readability, where the persons score is calculated by dividing the number of correctly identified words by the total number of blank spaces (Miller, 2009). The Cloze test has been shown to relate to the total number of blank spaces (Miller, 2009). The Cloze test has also been shown to produce significant and inverse correlations with readability formulas ($r = -0.581$). This is shown by using a Readability Analyser (RA) which implements a number of formulas such as F-K, Dale-Chall, Fry readability graph and others. This RA averages the RGL across the formulas used, therefore the inverse correlation suggests that as the RGL of a document increases, the readability of the document decreases (Gemoets et al., 2004). This is a simple and quick measure and can be used in conjunction with readability formulas to give more validity.

1.7 Russian language readability analysis

Readability measures for the Russian language have not been thoroughly researched. To date there are no specific Russian readability formulas. For this reason, the

readability analysis used formulas which are well known and are able to be adapted to use for different languages such as Russian.

1.7.1 Attributes of the Russian language

The Russian language is an East Slavic language which uses the Cyrillic script for the alphabet. The Russian alphabet is comprised of 33 letters. There are ten parts of speech of the Russian language: substantives which can be modified depending on their number, gender and case, adjectives, numerals, pronouns, verbs, participles, adverbs, conjunctions and interjections (Heard, 1827).

1.7.2 Russian language readability measures

There have not been any published readability formulas for the Russian language. A study conducted by Rock (1969) looked at developing a readability graph as a guideline for teachers of Russian to select appropriate reading materials for the students who were learning Russian in the US. A study conducted for a Russian PhD thesis by Oborneva (2006) as cited in Karpov et al. (2014), adapted the formulas of Flesch and Felsch-Kincaid for the Russian language. She did this by comparing the average length of sentences in both English and Russian words and the multi-syllable words in dictionaries for both languages and made adjustment coefficients (Karpov et al., 2014).

1.8 Research questions

The study aims and research questions are stated in detail below.

1.8.1 Study aims

The aim of this study was to describe the readability of hearing-related Internet information for the Russian language and to compare these results with the results of the Cloze test administered to native speakers of Russian.

1.8.1.1 Part 1

This part of the study aims to look at the following research questions:

(1) Is the readability of the top ten webpages from the six main Google domains when searching for information related to hearing in Russian higher than a RGL of six?

(2) Is there an even distribution in locality of webpages?

(3) Is there a significant difference in the mean RGL between webpages based on locality?

(4) Is there a significant positive relationship between the RGLs for each formula?

1.8.1.2 Part 2

This part of the study aims to compare results of two Cloze tests. The research question asked: Is the Cloze test with the lowest mean RGL easier to complete than the Cloze test with the highest mean RGL and is there a significant difference?

1.8.2 Study hypotheses

The hypotheses of this study are stated in detail below.

1.8.2.1 Part 1

(1) I hypothesise that the readability of the collected webpages will have a RGL higher than six.

(2) I hypothesise that there will not be an even distribution of locations that the webpages are from.

(3) I hypothesise that there will not be a significant difference in the mean RGL between webpages based on localities

(4) I hypothesise that there will be a significant positive relationship between the RGLs for each formula.

1.8.2.2 Part 2

I hypothesise that the Cloze test with the lowest mean RGL will be easier to complete than the Cloze test with the highest mean RGL and there will be a significant difference between the results.

2.0 Method

2.1 Part 1

The current study examined the readability of online hearing-related information available to the public in six countries where Russian is spoken by at least 5% of the population. These countries were: Russia, Kazakhstan, Ukraine, Belarus, Moldova, and Turkmenistan.

This study's methodology was based on a study conducted in 2012 by Laplante-Lévesque and colleagues who investigated the readability and quality of hearing related information on the internet for adults who have a hearing impairment and their significant others. This study was conducted in the English language, unlike the current study which will assess Internet information in the Russian language.

The readability assessment of the online hearing-related information consisted of analysing the first 10 Webpages retrieved using the Google country-coded Top-Level Domain (ccTLD) for each country, for the key search terms; "deterioration of hearing"- "Ухудшение слуха", "hearing impairment"- "Нарушения слуха", and "hearing loss"- "потеря слуха", and "снижение слуха". A ccTLD is defined as a unique two-letter sequence given to a particular country or geographical area to identify it, for example, .nz or .ru (Postel, 1994).

Readability was analysed using four readability formulas: Flesch-Kincaid (F-K; Kincaid, Fishburne, Rogers, & Chissom, 1975), Simple Measure of Gobbledygook (SMOG; McLaughlin, 1969), Dale-Chall readability (Dale & Chall, 1948), and Coleman-Liau Index (Coleman & Liau, 1975).

2.1.1 Search terms

The search terms used in this study were obtained from ten informants. The informants were native speakers of Russian and were either bilingual and lived in New Zealand or only spoke Russian and lived in Russia. The inclusion criteria of participants were: (1) they had to be native speakers of Russian, and (2) they could be of any age. Participants' knowledge of hearing health was not required. The informants were family friends as well as acquaintances from the Russian community in Christchurch, New Zealand. The search terms were collected by telephone. All participants were posed the same question

in Russian. The following was asked of participants over the phone. The English translation can be seen below, for the Russia version see Appendix A:

"I need some help with my thesis work and was hoping you could participate. I am looking at what words people use in Google search to look for information about anything hearing related. So, if I asked: If you wanted to find information on anything hearing related on the Internet, what would you search in Google?"

A limitation of this was that I did not specify that it was for a gradual hearing loss as some informants came up with terms for tinnitus and ear infection which was too specific and did not get included in the study.

Initially there were only three search terms that were acquired from informants. Google Trends (www.google.com/trends) was used to find any related search terms, the frequency of use of these terms and their location. Upon entering the search terms into Google Trends, no data could be found for the Russian search terms in any of the Russian-speaking Google Domains.

One informant commented that in Russian there are two terms for "hearing impairment" and these terms are interchangeable; where if someone used either term, any Russian speaker would understand the topic. An investigation into this new key word resulted in a better understanding of what the word actually means. In English as according to the Cambridge dictionary, "Ухудшение слуха" translates to "deterioration of hearing". Google trends were used again to make a comparison whether this word was used more frequently than the counterpart "hearing impairment", and it showed that they were used equally. This search term was then added. The total number of search terms used in this study was four.

One informant had said they would search in a question form, however the same identified key words were used in the question therefore the question itself was not included. It has been suggested to keep Google searches simple and to use important words, for example starting with a main keyword, if the results are not what the person is looking for then you can gradually add more words (Hindy, n.d.). When searching with a question, it can narrow the results down of the search because you are using more words and could potentially be being more specific. As this study had no specificity in the type of hearing-related information that would be searched, it was not relevant to have the question as a separate search term.

2.1.2 Search engines

To obtain the webpages for examination, six ccTLDs were used. These six ccTLDs were acquired by the research team who obtained information about where Russian was spoken as a first language by at least 5% of the population. They then looked into whether these countries had their own ccTLD as only the countries with a ccTLD could be used in this study. The six countries with their specific Google search engines were: Google Russia (Google.ru), Google Ukraine (Google.com.ua), Google Belarus (Google.by), Google Moldova (Google.md), Google Kazakhstan (Google.kz), and Google Turkmenistan (Google.tm). A search was then performed to identify the most popular search engine in each of these countries. According to StatCounter Global stats website, from April 2016 to April 2017, all six countries used Google as their dominant search engine (StatCounter Global Stats, 2017).

2.1.3 Internet search

The unit of analysis for the readability part of this study was a webpage. A webpage is defined as a smaller part of a larger website which has more specific information on a

particular topic (Christensson, 2006). A website is defined as a group of World Wide Web pages that contain hyperlinks to each other and are available online (Merriam-Webster, n.d.). Websites were not analysed in this study as webpages offer more samples for analysis. Twenty-four separate searches were completed (6 localities x 4 search terms). Duplicate webpages were defined as webpages whose contents had been previously identified by the search by either using a different search term or ccTLD and were excluded. This resulted in 46 unique webpages to include in the analysis. A discussion with the research team suggested that webpages that had information for children and their families that was not specifically hearing related (e.g. cochlear implants), or were purely medical (e.g. otitis media) should not be included. It was also agreed that pages that had information which could not be copied should be excluded. After applying the discussed recommendations, 39 unique webpages remained in the sample.

2.1.4 Inclusion/exclusion criteria

The first 10 webpages that result from the Google search of each of the ccTLDs using each of the identified search terms that matched the inclusion and exclusion criteria were selected for the readability analysis. Webpages were included if they contained information relating to hearing or hearing impairment, the information was in the target language (Russian), the information was available to the public, and if the webpage contained information about the organisation hosting the website. Webpages were excluded if they were a Google-identified advertisement, a video, a directory listing, and if the webpage was less than 100 words in length. For each search, the following information was collected: The Uniform Resource Locator (URL) of the website, the organisation related to the website (the location of the organisation hosting the website and the type of organisation), and whether the website had HON certification.

2.1.5 Search procedure

When performing the search, each of the search terms was entered into the six domains. Google Russia was used first, then Belarus, Kazakhstan, Moldova, Ukraine, and Turkmenistan. The reason for having this order was because of the internet penetration rate which refers to how often people from each country have access to the internet, with Russia having the most. The first 10 webpages that met the inclusion and exclusion criteria for each of the search terms in each domain were saved. For each search, information about the webpage was recorded in an Excel file. This information included: the URL, the organisation hosting the webpage, the location of the webpage, whether there is HON certification, and the date the webpage was last updated.

The type of organisation was classified as either being commercial, non-profit or government in origin. The origin of the website was determined by searching who owned the website on the webpage in the section usually headed “О сайте” which translates to “about the site”.

To determine whether a website had HON certification, firstly, the website itself was searched to see if the HON code seal was displayed, then a Google extension for the HON code was downloaded which showed if a particular website had HON certification without having to search the entire website. A downfall with this extension was that if a particular website did not display that they had HON certification, it would not be identified when doing the searches. Therefore, the URL of each website was pasted into a section of the HON website which shows if there is HON certification or not.

The location of each website was determined by investigating the website itself. Another Google extension for Internet Protocol (IP) addresses of websites was downloaded

and helped in determining whether the initial location found on the actual website was correct or not.

The webpages themselves were then analysed to make sure the information was related to hearing impairment. The relevant information of each webpage was then copied and pasted into individual Word documents for the readability analysis. Information on the webpage such as pictures and references were not copied into the Word documents.

2.1.6 Readability measures

F-K, SMOG, Dale-Chall, and Coleman-Liau index were the specific formulas used in this study through the Russian readability website. This study's dependant variable was the mean readability score or RGL.

The F-K formula is a modified version of the Flesch Reading Ease formula developed by Rudolf Flesch in 1948 which gave a scoring from 0-100; where a score of 100 is very easy to read and 0 is very difficult (Dubay, 2004). The F-K produces a US reading-grade score; where 4^{th} grade and below is very easy to read and above 8^{th} grade is fairly difficult (Dubay, 2004).

The SMOG formula has been widely used for calculating readability of health-related information as it is one of the easiest to calculate (Ley & Florio, 1996). The SMOG formula is used by giving a RGL by first counting the number of polysyllabic words (words with 3 or more syllables) in three samples of ten sentences, then estimating the square root of the nearest perfect square and adding three to this square root (McLaughlin, 1969).

The Dale-Chall readability formula uses a list of 3000 words and the average number of words in a sentence to predict the mean grade level of people who are able to answer 50% of questions relating to a passage (Ley & Florio, 1996). According to a study performed by Wang and colleagues in 2013, 155 articles that used readability formulas were evaluated and found that the Dale-Chall readability formula was used less commonly, only 2.58% of the

time compared to the F-K which was used 57.42% and SMOG which was used 25.81% of the time.

Coleman-Liau index is the fourth readability formula that was used. This formula determines the grade level of the text based on the sentence length and character count (Coleman & Liau, 1975).

All above readability formulas and how to calculate them can be viewed in Appendix B.

2.1.7 Readability analysis

The readability analysis was performed using a Russian readability website: http://ru.readability.io/. This website allows members of the public to perform a readability analysis on a particular text or URL. The website does this by using multiple readability formulas on the particular text that is inserted. As well as providing an analysis on different readability formulas, the website provides details of how many words are in the text, how many syllables they have, how many sentences there are, the number of polysyllabic words, and the percentage of compound words.

The analysis was performed by the following procedure:

1. 'Text' option was chosen instead of 'URL' so the full copied text could be pasted.
2. The information of each word document containing the information from the webpages was copied and pasted into the Russian readability website.
3. F-K, Coleman-Liau index, Dale-Chall, and SMOG scores were given automatically.
4. The readability scores for each formula were recorded in a Microsoft Excel spreadsheet.
5. This procedure was repeated for all saved Word documents of all webpages collected.

2.1.8 Planned statistical analysis

The statistical analyses for this study was A Univariate analysis of variance (ANOVA) to test for differences in the mean RGL between webpages based on locality. A Chi-Squared test was also used to analyse the distribution of locality of the webpage. A one-tailed Pearson correlation was used to analyse the statistical significance of the RGLs derived for each formula.

2.2 Part 2

The Cloze test was administered to ten participants. This test was performed to make a comparison between the webpage with the highest and lowest readability levels that were found in part 1.

2.2.1 Participants

A low risk human ethics application was sent into the Human Ethics Committee at the University of Canterbury for the approval of requiring participants for this study. The ethics committee reviewed and approved the application. The approval letter can be seen in Appendix C. Ten native Russian speakers who have no expertise in hearing health provided data for the Cloze test in this study. These participants were selected as they were used as informants for the search terms in Part 1 of this study. The participants lived in both Russia and New Zealand. The age of participants was between 43-54 years. There were two male participants and eight female participants. The participants who lived in New Zealand had been living here for 15-20 years and originally came from the Far East of Russia, Kazakhstan and Ukraine. The participants who live in Russia were from the Far East of Russia.

2.2.2 Procedure

When performing Part 2 of the study, two Cloze tests were produced. The first Cloze test was derived with a sample of text from the website in Part 1 which had the lowest mean readability score. Every fifth word of the 262-word text was deleted. The second Cloze test was derived with a sample of text from the website that had the highest mean readability score in Part 1. Again, every fifth word of the 254-word text was deleted. Once the tests were developed, the participants used in part 1 were contacted and asked to complete the Cloze tests. An email was sent to them which mentioned the topic of the study and requesting them to email back if they would like to participate.

The participants who responded were then sent an information sheet and the study materials explaining the entire study and what they needed to do. Initially, ten participants were contacted; all barring one participant completed both Cloze tests. A replacement participant was contacted and they completed the Cloze tests, however a communication issue arose leading to an unreliable result for one of the Cloze tests which resulted in this participant's data not being used. Another potential participant was contacted and they agreed to complete the tests. Upon commencement of this study, ten participants were used. The replacement participants were found through personal contacts.

2.2.3 Cloze test calculation

The percentage correct for both Cloze tests was calculated by dividing the number of correct responses by the total number of blanks then multiplying by 100. This was completed for each participant for both Cloze tests. SPSS was then used to calculate descriptive statistics.

2.2.4 Planned statistical analysis

A non-parametric related samples test (Wilcoxon signed ranks test) was used to analyse if there was a significant difference in the mean scores between the two Cloze tests.

3.0 Results

3.1 Overview

In total, 39 webpages were analysed in the current study to identify the readability of hearing-related information available online in Russian. The analysis required the top 10 results for the search terms “deterioration of hearing”, “hearing loss”, and “hearing impairment” in Russian. This resulted in four search terms when translated into Russian. The search was completed using six Google ccTLDs which corresponded to the six main localities where Russian is spoken by at least 5% of the population. Each country (Russia, Ukraine, Belarus, Moldova, Kazakhstan, and Turkmenistan) uses Google as their dominant search engine (StatCounter Global Stats, 2017). During the retrieval process of the webpages, duplicates were not included. Webpages that had medical information about hearing loss, such as otitis media were also excluded as they were specific only to the management of that one problem.

The Cloze procedure was used to validate the results of the readability scores obtained through the use of the formulas. Ten participants who are native speakers of Russian were used for this part of the study; these were the same people who were the informants for the search terms in part 1 of the study. Two Cloze tests were developed from the highest and lowest readability scoring webpages. Participants were required to complete both Cloze tests so a comparison could be made between the scores.

3.2 Part 1

3.2.1 Description of webpages

Of the 39 webpages analysed, 35 were of a commercial organisation, two were non-profit and two were government. The analysis of whether there was an even distribution in the type of organisation across the websites did not take place as it was apparent that there was not an even spread among the three types of organisations. Of the six localities that the websites were retrieved from, there were only three main locations that majority of the webpages were from. These locations were Russia, Ukraine and World. The location "world" refers to the websites where a person is able to choose a specific country and the information of the website becomes country and language specific. There were 18 webpages of Russian origin, 10 were from Ukraine, 6 were counted as world and the other 3 were from Belarus, Latvia and Australia. Belarus, Latvia and Australia were not included in the analysis as there was only one webpage from each. Another aspect of the webpage that was examined was whether the webpage had the HON code. Of the 39 webpages, only two had the HON code which meant this too could not be analysed as there was no variability. The last aspect that was examined was the readability scores of the four readability formulas. The results of the minimum, maximum and mean readability scores are shown in table 1 below.

Table 1. Results of the lowest, highest, and mean scores for each readability formula.

Formula	*Minimum RGL*	*Maximum RGL*	*Mean RGL*
F-K	9.11	24.66	16.44
Dale-Chall	7.41	19.01	13.02
Coleman-Liau	8.99	21.96	14.80
SMOG	8.48	21.16	13.56

Mean RGL	8.50	20.54	14.46

Note: RGL = Reading grade level, F-K = Flesch-Kincaid, SMOG = Simple Measure Of Gobbledygook

The Lowest readability score was given by the Dale-Chall formula of 7.41, and the highest by the F-K formula of 24.66. The lowest mean readability score was given to the Phonak website: https://www.phonak.com/ru/ru/poterya-sluha.html which had a RGL of 8.50. The website with the highest mean readability score was: http://www.deafworld.ru/poleznaj_info/Vidi_naruheniy_sluha which had a RGL of 20.54. For a full list of the webpages and their RGLs refer to Appendix D.

For all four analysed readability formulas there was no skewness of the data, meaning it was evenly distributed. Two outliers were found in the data; however they were not significant meaning parametric testing could be performed.

3.2.2 Results of hypothesis testing

The readability of webpages was analysed using four readability formulas; F-K, Dale-Chall, Coleman-Liau and SMOG. To answer the first question - Is the readability of the first ten webpages from the six main Google ccTLDs when searching for information related to hearing in Russian higher than a RGL of six - a one-sample t-test was used to demonstrate whether or not there was any statistical significance, results can be seen in the table below.

Table 2. Table of one-sample T-Test results.

	Mean	*SD*	*t*	*df*	*P value*	*Mean difference*	*95% CI*
Mean RGL	14.46	2.76	32.75	38	< .000	14.46	13.57-15.35

Note: RGL = Reading grade level, SD = Standard deviation, t = t-statistic, df = Degrees of freedom, P value = Pearson correlation, CI = Confidence interval.

The results show that the null hypothesis was not supported. This means that the readability scores had a mean RGL significantly higher than six.

A non-parametric test (chi-square) was performed to see if there was an even distribution between localities. The results showed that the null hypothesis was supported: there was an even distribution of webpage country of origin, χ^2 (2, N = 36) = 4.67, p = .097.

A one-way ANOVA was used to investigate if there was a significant difference in the mean RGL between webpages based on locality. The table below describes the mean RGLs for each of the three locations.

Table 3. Table of mean RGLs based on location

Location	*Mean*	*SD*
Russia	14.48	2.28
Ukraine	14.92	3.26
World	14.50	3.52
Total	14.60	2.79

The table above shows that the mean RGLs for the three localities are all very similar. The Levene's statistic found that the assumption of homogeneity of variance was met, p = .112. The ANOVA supported the null hypothesis that the RGLs between localities were not significantly different; ($F(2,39) = .082$, $p = .921$, $\eta^2 = .005$).

The final analysis answered the question; is there a significant positive relationship between the RGLs for each formula? A bi-variate correlation was run and a table of results can be seen below.

Table 4. Table of Pearson correlation.

	F-K	*Dale-Chall*	*Coleman-Liau*	*SMOG*	*Mean RGL*
F-K	1	.956	.958	.885	.990
Dale-Chall		1	.893	.926	.981
Coleman-Liau			1	.772	.944
SMOG				1	.930
Mean RGL					1

The results showed that there were significant positive relationships for all correlations at the p = 0.01 level (2-tailed). All readability formulas were highly correlated as shown in the table above, meaning all analysed webpages had similar readability scores, no matter what formula was used.

3.3 Part 2

3.3.1 Participant Summary

A questionnaire was required to be filled in by participants upon completing the study. The participants were required to state their age, gender, ethnicity, and the number of years of schooling they had completed. This information was required to get a general idea about the participants and be able to make appropriate decisions when analysing the data in

case there were any inconsistencies. The participants' ages ranged between 43 and 54 years. There were eight females and two males who took the tests. Nine participants are of Russian ethnicity and one is Ukrainian. All ten participants had completed 14 years and over of schooling which included both school and tertiary education.

3.3.2 Results of hypothesis testing

The results showed that the Cloze test that was designed from the website with the lowest readability score was easier to understand than the Cloze test from the website with the highest readability score. The mean percentage for the Cloze test with the lowest readability score was 64.08%, while the mean percentage for the Cloze test with the highest readability score was 45.78%. The Wilcoxon signed ranks test was used to perform the analysis to see if there was any statistical significance in Cloze test scores between the two websites. This test was used as it compares two sets of scores that come from the same participants. It is a non-parametric test as the distribution of the participant scores was not normal. The results indicated the Cloze test with the low readability score was significantly easier to complete (M = 64.08%) than the Cloze test with the high readability score (M = 45.78%), $z = -2.40$, $p = .007$, $r = 0.61$.

3.4 Results summary

In conclusion, all analysed webpages had higher readability levels than the optimal level of six. The general public would need at least 13-16 years of education to understand and comprehend the information available on these webpages. No significant differences were found for the mean readability levels of webpages from different origins as all readability levels were very similar. The Cloze test concluded that the webpage with the highest readability score was more difficult to complete than the Cloze test with the lowest readability score

4.0 Discussion

4.1 Overview

The aim of this study was to describe the readability of online hearing-related information in the Russian language and to compare the readability results with the results of the Cloze procedure that was administered to native speakers of Russian. Five hypotheses were formulated for this study. These were: 1) that the readability of the collected webpages will have a RGL higher than six. The null hypothesis was not supported as all the analysed webpages had an RGL higher than six. 2) There will not be an even distribution of where the webpages are located. This null hypothesis was supported as there was an even distribution of webpage country of origin. 3) There will not be a significant difference in the mean RGL between webpages based on localities. The null hypothesis was supported as the RGLs between localities were not significantly different. 4) There will be a significant positive relationship between the RGLs for each formula. The null hypothesis was not supported. All readability formulas were highly and significantly correlated which showed that all webpages had similar readability scores no matter what formula was used. 5) The Cloze test with the lowest mean RGL will be easier to complete than the cloze test with the highest mean RGL and there will be a significant difference between the results. Results showed that the Cloze test with the lowest mean readability score was significantly easier to complete than the cloze test with the highest mean readability score.

The current study analysed 39 unique webpages in Google ccTLDs where Russian is spoken by at least 5% of the population. There is no published study on the readability of hearing-related Internet information in Russian. To date there has also been limited information available on the readability and quality of online information in languages other than English. Six Google ccTLDs were used to extract the appropriate webpages. Upon

retrieval, each webpage was examined for information such as the organisation hosting the webpage, the location, whether the webpage had HON certification and the date the webpage was last updated. However, not all websites displayed some of this information and direct comparisons could not be made. The following sections will discuss some previous literature on the topic of readability, the results of the readability and cloze tests of the analysed websites as well as the validation of the HON code of these websites

4.2 Previous literature

Readability of online health information in English has extensively been researched over the years. It is important that the information that is provided to health professionals, patients and the general public is at a level where it is easy to read and understand. More and more literature has also been emerging on the topic of readability in other languages. Sentence length and vocabulary difficulty were used to predict the difficulty level of a text from the 1920s (Dubay, 2004). By the 1980s, there were thousands of published studies on the validity of readability formulas, as well as around 200 formulas available for use (Dubay, 2004). It has been assumed that adult patients are able to read at levels that equal to the years of schooling they had completed (Doak, Doak, & Root, 1996). However, this assumption is incorrect in most cases. Literacy skills have been shown to contribute a great deal on an individual's RGL and that individual's literacy skills tend to be around five grade levels lower than the last year of school completion (Doak, Doak, & Root, 1996). This can cause a number of problems when individuals are searching for health care information online.

Having limited Internet literacy skills, the inability to navigate the Internet, the inability to judge if the information they find is reliable and also inadequate health literacy skills in general to be able to put the information they found to use are all issues that consumers face when searching online (Atcherson et al., 2014). To facilitate this health literacy problem, it has been suggested that health materials for adults are written at the

recommended 5th-6th RGL (Doak, Doak, & Root, 1996; Dubay, 2004). Unfortunately, throughout this study it has been found that this recommended RGL has not been applied to online health information in English, Chinese and now Russian even when using different formulas.

Comprehension tests such as the well-known Cloze test developed by Taylor in 1953, has also been widely used as a predictor of an individual's understanding of a text. This test is based on the theory that as an individual's reading level improves; they are better able to fill in the blanks of the test (Dubay, 2004). Although this may true, it has been shown that readers are less mindful of the difficulty of the article when they are interested in the topic resulting in better comprehension of the text (Pauk, 1973; Belloni & Jongsma, 1978). The opposite can be said about readers who are uninterested; they are less likely to score well on comprehension tests as they are unmotivated to fill in the blanks (Belloni & Jongsma, 1978).

Health care information can be very difficult to read and understand especially when individuals are not familiar with the jargon used. A study conducted by Ritchie et al. (2015), looked at online information available to the public on glue ear. They found that the information specific to glue ear had significant variation between the websites that were analysed. Of the 27 websites the mean F-K score found was 49.7 with a range of 25.8-65.7 which corresponds to a text that is very difficult to read and a text that is standard/uses plain English respectively (Flesch, 1948). Although there was a large variation in the scores, no website was found to be at the recommended 5-6 RGL. The score of 65.7 corresponded to 8-9 RGL which was the lowest (Flesch, 1948). Another study conducted by Atcherson et al (2014) analysed the readability of materials on the American Speech-Language-Hearing Association (ASHA) website which can be used by consumers to look at information relating to audiology and speech language pathology. The study concluded that of the 225 documents analysed, 85.4% of them had a RGL higher than the recommended RGL of 5-6. This result

demonstrates that consumers who have low literacy will most likely misinterpret the information they read or will not be able to make informed decisions on what to do with their own or their family's hearing health problems. For this reason it is important for health care materials, especially hearing health materials to be written at an appropriate level.

According to the Global Burden of Disease Project, it is estimated that hearing loss is the third leading cause of disability (Stevens et al., 2013). There has been a global increase of the prevalence of hearing loss over the years and this is why it is so important that information available to the public is comprehendible. Unfortunately until now no data has been available in the Russian language of online hearing-related information. This is a concern as Russian is commonly spoken all over the world. Fortunately this study gives insight into this issue and more information in other languages can be analysed in the future.

4.2.1 Readability formulas

The analysed webpages were found to contain information with a higher than recommended readability level. The F-K formula was found to have the highest mean RGL of 16.44, followed by Coleman-Liau of 14.80, SMOG followed with 13.56 and the lowest mean RGL was from the Dale-Chall formula of 13.02. The hearing-related health information from the analysed webpages was shown to be significantly higher than the recommended RGL of 5-6. The results show that individuals require approximately 13-16 years of education to be able to read and comprehend the information on these webpages. This result meant that the null hypothesis could not be supported.

This can have implications for consumers as they are unable to make appropriate health decisions due to the difficulty of what they are reading online, however, the help of health care professionals, and being able to provide simple and concise healthcare information can lead to positive health outcomes for the consumer. Clinicians can have a

positive impact on healthcare information by making sure their healthcare materials are at a RGL that is appropriate for consumers. Clinicians also need to stay informed of the ever-changing research available so they can direct clients to appropriate websites with the correct information.

No comparison could be made between the types of organisation that was hosting the webpage as there was not enough spread among the 39 webpages analysed. The majority of webpages were of commercial origin, two were non-profit and two were government.

Of the six locations that the webpages were collected from, only three could be compared as that was where majority of the webpages came from. These locations included Russia, Ukraine and World. The results concluded that there was no significant difference of RGLs between location meaning that no matter where the origin of the website was they were all difficult to read and comprehend.

As there were four readability formulas, a correlational analysis was used to see if the formulas were significantly related. The results showed that there were significant positive relationships of all readability formulas concluding that all formulas produced similar readability scores. These results are on trend with the literature previously discussed. This provides evidence of the validity of the readability formulas as they are all producing high RGLs. From this study, it has been shown that readability formulas are a quick and easy way to ensure healthcare materials online are at an appropriate level for consumers so they are able to make the best health choices for themselves.

4.2.2 HON code

The HON code is a code of conduct that a health care website should have as it helps create a standard of reliability of the medical and health information available on the Internet to patients, professionals and the public (HON, 2017). Websites that abide with the HON code require applying eight key principals that make sure the website is up to standard with

information that is useful, reliable and not misleading. It has been recorded that in August 2010, there were approximately 7300 websites that had the HON certification which is a small proportion of health information material that is actually on the Internet (Laplante-Lévesque et al, 2012a). The study conducted by Laplante- Lévesque and collegues (2012) looked at the quality and readability of Internet information in English for adults and their significant others with a hearing impairment. This study assessed 66 websites by looking at the origin of the website, date of last update, quality, and readability. They found that of the 66 websites only 9 had the HON certification. Of these 9 websites, the ones that were of government origin were found to be more likely to have HON certification.

When the quality and readability of websites were compared, the authors found that readability was independent of whether or not the website had the HON code. This indicates readability was still high regardless of any other factors that had been looked at such as origin, date of last update and quality (Laplante- Lévesque et al., 2012a). Of the 39 webpages analysed in the current study, only two had the HON code. The webpages with the HON certification in this study were of commercial and government origin. They both displayed a date of last update which is one of the HON requirements (HON, 2017). However, these two webpages did not have the lowest average readability scores even when following the HON code principals. This is consistent with the findings in the above study which means that even with certification, websites are still hard to read and comprehend.

4.2.3 Cloze test

The Cloze test is a tool used for measuring readability (Taylor, 1953). This test has been validated against readability formulas such as the Fry and Dale-Chall using literary passages (Taylor, 1958). A study conducted by Gemoets and colleagues (2004), assessed the readability of health materials for consumers by using readability formulas as well as the

Cloze test using actual readers. They used 20 consumer health documents for the study and found that as the readers were scoring higher on the Cloze test, the readability scores were decreasing, showing a significant correlation between the Cloze test and readability scores. The current study also found this pattern. Two of the analysed webpages of this study were used for the Cloze test. The webpages used were the ones that had the highest and lowest readability scores derived from the readability formulas. Both Cloze tests were completed by ten participants so a comparison could be made between the scores. The results concluded that the Cloze test with the low readability score was significantly easier to complete than the Cloze test with the high readability score. The mean percentage of correct answers for the low readability Cloze test was around 64% and the high readability Cloze test was around 45%. Although it is clear that the low readability Cloze test was easier than the other test, participants still had trouble completing the lower Cloze test. This concludes that the readability levels of all the analysed webpages are still too high especially if the person looking for the information knows nothing about hearing impairment

4.3 Study limitations

The present study encountered some limitations throughout the search process. The informants used in this study were all family friends as this was the quickest and easiest way to get the required information. This could have been a selection bias as they may have felt obliged to help or the opposite and felt like they did not have to put much effort in. The informants were required to complete two parts of this study and the second part was going to take longer than the first. Some informants were spouses and the test was required to be completed by both parties individually. This may not have been clear to some as the completed tests only came back from one individual (either just the wife or husband). This loss in translation meant it took longer to get results back and also resulted in someone not

completing it at all as it took too long. If the informants were not people that were known to me or my family, the completion rate as well as quality of the answers may have been better.

Another limitation was the way the question was posed to the informants. The question for the search terms were asked over the phone. I did not specify that the information being looked at was for a gradual hearing loss leading informants to come up with specific search terms such as “tinnitus” and “ear infections”. These search terms were not included as they were too specific for the nature of this study. It may have also been easier to have participants come and do the Cloze tests in a controlled environment where they could not turn to the Internet or any other sources for help. This could have limited any confounding factors. A further limitation of this study was the age range of participants. Although there were only ten, there could have been a larger range as younger or older people may be using different search terms for this topic.

4.4 Clinical implications

Readability of the Russian language has not been well documented, and currently there have not been any studies looking at the readability of hearing-related information on the Internet of different countries in the Russian language. As it has been touched on, many individuals all over the world turn to the Internet for hearing-related health information. It is crucial that this information is at a level that is understood by the Russian public so informed decisions about their hearing health can be made. The results of this study showed that the hearing-related information on the Internet for the Russian public is unsatisfactory as the readability is at levels which are too difficult for most people to comprehend. Clinicians and health care professionals need to make sure they are giving their patients high quality material that is easy enough to understand so that greater client satisfaction can be achieved.

Research has shown that healthcare providers are still a trusted and reliable source of information and they are able to influence a person's decision more so than any other source of information (Pennbridge, Moya, & Rodrigues, 1999; Hesse et al., 2005). This is important as it can strengthen the bond between the healthcare provider and patient. The use of the Internet to search for healthcare materials is only going to increase as more people gain access to the Internet. It is important clinicians are aware of how to interact with patients who access online healthcare materials. Patient's coming in with their own background knowledge from online materials has already been documented. This can hinder the patient-healthcare provider relationship as not every treatment plan or medication is suitable for everyone and some patients may not understand that and still request to be treated with what they read (Heese et al., 2005). Clinicians need to be able to take the information the patient gives and work with them collaboratively to find the best possible outcome for the patient.

4.5 Future research

The present study gives a first look into the readability of online hearing-related information in Russian. The next step would be to delve deeper and look into how the public search for hearing-related information specifically. This study used only ten informants/participants. If there were more informants for collecting search terms there may have been a bigger variety and therefore more webpages to analyse. It would also be important to see how people decide on whether information is reliable or not and how many different sources they go to before making a decision. It would be a good idea to conduct this study in the countries where Russian is spoken by at least 5% of the population or actively seek out people from each country to participate. This would give more of an indication as to whether the search terms would be the same as what was acquired in this study or not.

4.6 Conclusion

This study examined the readability of online hearing-related information available to Russian consumers. The readability of the webpages was analysed using four formulas which included the F-K, SMOG, Dale-Chall and Coleman-Liau. The Cloze test was also used to compare the highest and lowest readability scoring webpages. In conclusion, the readability of all 39 webpages analysed were higher than the recommended RGL of 5-6. The Cloze test scores showed that the webpage with the lowest readability score was easier for participants to fill in than the Cloze test with the highest readability score. This was expected, although the lower Cloze test was still difficult to complete. These results suggest that people may not be able to make informed decisions about their own or their families hearing health as the information available to them is too difficult to read and comprehend. However, clinicians can have a role in helping patients make informative decisions by providing appropriate resources and information and being up to date with current research. Making sure the company websites are at a standard that is suitable for most people to understand is also important as it would result in greater hearing health outcomes for the patients.

References

Atcherson, S. R., DeLaune, A. E., Hadden, K., Zraick, R. I., Kelly-Campbell, R. J., & Minaya, C. P. (2014). A computer-based readability analysis of consumer materials on the American Speech-Language-Hearing Association website. Contemporary Issues in Communication Science and Disorders, 41(Apr), 12-23.

Agrawal, Y., Platz, E. A., & Niparko, J. K. (2008). Prevalence of hearing loss and differences by demographic characteristics among US adults: Data from the national health and nutrition examination survey, 1999-2004. Archives of Internal Medicine, 168(14), 1522.

Batterham, R. W., Hawkins, M., Collins, P. A., Buchbinder, R., & Osborne, R. H. (2016). Health literacy: Applying current concepts to improve health services and reduce health inequalities. Public Health, 132, 3-12. doi:10.1016/j.puhe.2016.01.001

Baker, L., Wagner, T. H., Singer, S., & Bundorf, M. K. (2003). Use of the internet and E-mail for health care information: Results from a national survey. Jama, 289(18), 2400-2406. doi:10.1001/jama.289.18.2400

Belloni, L. F., & Jongsma, E. A. (1978). The effects of interest on reading comprehension of low-achieving students. *Journal of Reading*, *22*(2), 106-109.

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low Health Literacy and Health Outcomes: An Updated Systematic Review. Annals of Internal Medicine, 155(2), 97. doi:10.7326/0003-4819-155-2-201107190-00005

Bormuth, J. (1969). Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability. Reading Research Quarterly, 4(3), 358-365. doi:10.2307/747144

Brashers, D. E., Goldsmith, D. J., & Hsteh, E. (2002). Information seeking and avoiding in health contexts. *Human Communication Research, 28*(2), 258-271. doi:10.1093/hcr/28.2.258

Cambridge Dictionary | Free English Dictionary, Translations and Thesaurus. (n.d.). Retrieved from http://dictionary.cambridge.org/

Chen, H. L. (1994). Hearing in the elderly. relation of hearing loss, loneliness, and self-esteem. Journal of Gerontological Nursing, 20(6), 22.

Chisolm, T. H., Johnson, C. E., Danhauer, J. L., Portz, L. J. P., Abrams, H. B., Lesner, S., Newman, C. W. (2007). A Systematic Review of Health-Related Quality of Life and Hearing Aids: Final Report of the American Academy of Audiology Task Force on the Health-Related Quality of Life Benefits of Amplification in Adults. Journal of the American Academy of Audiology, 18, 151-183.

Christensson, P. (2006). Web Page Definition. Retrieved 2017, Jun 22, from https://techterms.com

Cline, R. J. W, Haynes, K. M.; Consumer health information seeking on the Internet: the state of the art, Health Education Research, Volume 16, Issue 6, 1 December 2001, Pages 671–692, https://doi-org.ezproxy.canterbury.ac.nz/10.1093/her/16.6.671

Coleman, M., & Liau, T. L. (1975). A computer readability formula designed for machine scoring. Journal of Applied Psychology, 60(2), 283-284. doi:10.1037/h0076540

Collins-Thompson, K. (2014). Computational assessment of text readability: A survey of current and future research. ITL - International Journal of Applied Linguistics, 165(2), 97-135. doi:10.1075/itl.165.2.01col

Cotten, S. R., & Gupta, S. S. (2004). Characteristics of online and offline health information seekers and factors that discriminate between them. Social Science & Medicine, 59(9), 1795-1806. doi:10.1016/j.socscimed.2004.02.020

Dale, E., & Chall, J. S. (1948). A formula for predicting readability: Instructions. Educational Research Bulletin, 27(2), 37-54.

Dalton, D. S., Cruickshanks, K. J., Barbara E K Klein, & Klein, R. (2003). The impact of hearing loss on quality of life in older adults. The Gerontologist, 43(5), 661-668. doi:10.1093/geront/43.5.661

Deering, M. J., & Harris, J. (1996). Consumer health information demand and delivery: Implications for libraries. Bulletin of the Medical Library Association, 84(2), 209-216.

Demographics of Internet and Home Broadband Usage in the United States | Pew Research Center. (2017). Retrieved from http://www.pewinternet.org/fact-sheet/internet-broadband/

Dillon, H. (2012). Hearing aids (2nd ed.). Sydney: Boomerang Press.

Doak, C., Doak, L., & Root, J. (1996). Teaching patients with low literacy skills (2nd ed.). Philadelphia:J.B. Lippincott Company.

DuBay, W. H. (2004). The Principles of Readability. Online Submission.

Dutta-Bergman, M. J. (2004). Primary sources of health information: Comparisons in the domain of health attitudes, health cognitions, and health behaviors. *Health Communication, 16*(3), 273-288. doi:10.1207/S15327027HC1603_1

Feder, K., Michaud, D., Ramage-Morin, P., McNamee, J., & Beauregard, Y. (2015). Prevalence of hearing loss among canadians aged 20 to 79: Audiometric results from the 2012/2013 canadian health measures survey. Health Reports, 26(7), 18.

Fisch, L. (1973). Epidemiology of congenital hearing loss. Audiology : Official Organ of the International Society of Audiology, 12(5), 411-425.

Flesch, R. (1948). A new readability yardstick. Journal of Applied Psychology, 32(3), 221-233. doi:10.1037/h0057532

Friedman, D. B., Hoffman-Goetz, L., & Arocha, J. F. (2006). Health literacy and the world wide web: Comparing the readability of leading incident cancers on the internet. Informatics for Health and Social Care, 31(1), 67-87. doi:10.1080/14639230600628427

Fry, E.B., J. E. Kress, and D. L. Fountoukidis. 1993. The reading teacher's book of lists: Third edition. West Nyack, NY: The Center for Applied Research in Education.

Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016. (n.d.). Retrieved from https://www.gartner.com/newsroom/id/3598917

Gemoets, D., Rosemblat, G., Tse, T., & Logan, R. (2004). Assessing Readability of Consumer Health Information : An Exploratory Study. Medinfo, 11(Pt 2), 869–873

Hesse, B. W., Nelson, D. E., Kreps, G. L., Croyle, R. T., Arora, N. K., Rimer, B. K., & Viswanath, K. (2005). Trust and sources of health information. the impact of the internet and its implications for health care providers: Findings from the first health information national trends survey. Archives of Internal Medicine, 165(22), 2618-2624. doi:10.1001/archinte.165.22.2618

Hindy, n.d. 20 Tips To Use Google Search Efficiently. [Blog post] (n.d.). Retrieved from http://www.lifehack.org/articles/technology/20-tips-use-google-search-efficiently.html

Health literacy | Ministry of Health NZ. (n.d.). Retrieved from http://www.health.govt.nz/our-work/making-services-better-users/health-literacy

Heard, J. (1827). A practical grammar of the Russian language. St Petersburg: Taylor institution.

Household Use of Information and Communication Technology: 2012. (n.d.). Retrieved from

http://www.stats.govt.nz/browse_for_stats/industry_sectors/information_technology_and_communications/HouseholdUseofICT_HOTP2012/Commentary.aspx

Introduction to Russian Language - Basic Facts. (n.d.). Retrieved from http://learningrussian.net/russianlanguage/

Johnson, W. 1946. People in Quandaries. New York: Harpers, Appendix

Karpov, N., Baranova, J., & Vitugin, F. (2014). Single-Sentence Readability Prediction in Russian. Communications in Computer and Information Science, 91-100. doi:10.1007/978-3-319-12580-0_9

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. Scientific Studies of Reading, 12(3), 281-300. doi:10.1080/10888430802132279

Kelly, L. J., Kahn, A. (1991) Illiteracy and hearing loss management: The readability of clinic forms. Journal of the Academy of Rehabilitative Audiology, Vol 24, 35-42.

Kelly, L. J. (1996). The Readability of Hearing Aid Brochures. Journal of the Academy of Rehabilitative Audiology, 24, 41-47.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel (No. RBR-8-75). Naval Technical Training Command Millington TN Research Branch.

Laplante-Lévesque, A., Knudsen, L. V., Preminger, J. E., Jones, L., Nielsen, C., Öberg, M., . . . Östergötlands Läns Landsting. (2012). Hearing help-seeking and rehabilitation: Perspectives

of adults with hearing impairment. International Journal of Audiology, 51(2), 93-102. doi:10.3109/14992027.2011.606284

Laplante-Lévesque, A., Brännström, K. J., Andersson, G., Lunner, T., Linköpings universitet, Institutionen för beteendevetenskap och lärande, Filosofiska fakulteten. (2012a). Quality and readability of english-language internet information for adults with hearing impairment and their significant others. International Journal of Audiology, 51(8), 618-626. doi:10.3109/14992027.2012.684406

Laplante-Lévesque, A., Thorén, E. S., Linköpings universitet, Handikappvetenskap, Institutionen för beteendevetenskap och lärande, & Filosofiska fakulteten. (2015). Readability of internet information on hearing: Systematic literature review. American Journal of Audiology, 24(3), 284.

Ley, P., & Florio, T. (1996). The use of readability formulas in health care. Psychology, Health & Medicine, 1(1), 7-28. doi:10.1080/13548509608400003

Love, R. L., & Bird, P. (2013). Cochlear implantation in mitochondrial deafness due to A7445G mutation. Cochlear Implants International, 14(1), 28-31. doi:10.1179/1754762811Y.0000000030

Madell, J. R., & Flexer, C. A. (2014). Pediatric audiology: Diagnosis, technology, and management (2nd ed.). New York, NY: Thieme.

Mathers, C., Smith, A., Concha, M. (2003). Global burden of hearing loss in the year 2000. Genava: World Health Organization.

Mc Laughlin G. H. (1969). SMOG grading-a new readability formula. Journal of Reading, 12(8), 639-646.

McCallum, D. R., & Peterson, J. L. (1982). Computer-based readability indexes. Proceedings of the ACM '82 conference on - ACM 82. doi:10.1145/800174.809754

Mcinnes, N., & Haglund, B. J. A. (2011). Readability of online health information: Implications for health literacy. *Informatics for Health and Social Care*, *36*(4), 173-189. doi:10.3109/17538157.2010.542529

McMullan, M. (2006). Patients using the internet to obtain health information: How this affects the patient–health professional relationship. Patient Education and Counseling, 63(1), 24-28. doi:10.1016/j.pec.2005.10.006

Metzger, M. J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. Journal of the American Society for Information Science and Technology, 58(13), 2078-2091. doi:10.1002/asi.20672

Miller, M. J., DeWitt, J. E., McCleeary, E. M., & O'Keefe, K. J. (2009). Application of the cloze procedure to evaluate comprehension and demonstrate rewriting of pharmacy educational materials. The Annals of Pharmacotherapy, 43(4), 650–7

Monzani, D., Galeazzi, G. M., Genovese, E., Marrara, A., & Martini, A. (2008). Psychological profile and social behaviour of working adults with mild or moderate hearing loss. Acta Otorhinolaryngologica Italica : Organo Ufficiale Della Società Italiana Di Otorinolaringologia e Chirurgia Cervico-Facciale, 28(2), 61.

Morahan-Martin, J, M. How Internet Users Find, Evaluate, and Use Online Health Information: A Cross-Cultural Review CyberPsychology & Behavior. November 2004, 7(5): 497 510.https://doi.org/10.1089/cpb.2004.7.497

Niskar, A. S., Kieszak, S. M., Holmes, A., Esteban, E., Rubin, C., & Brody, D. J. (1998). Prevalence of hearing loss among children 6 to 19 years of age: The third national health and nutrition examination survey. Jama, 279(14), 1071-1075. doi:10.1001/jama.279.14.1071

Nondahl, D. M., Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, R., & Klein, B. E. (1998). Accuracy of self-reported hearing loss. Audiology : Official Organ of the International Society of Audiology, 37(5), 295-301.

Nutbeam, D. (2008). The evolving concept of health literacy. *Social Science & Medicine, 67*(12), 2072-2078. doi:10.1016/j.socscimed.2008.09.050

Online Activities in Emerging and Developing Nations | Pew Research Center. (n.d.). Retrieved from http://www.pewglobal.org/2015/03/19/2-online-activities-in-emerging-and-developing-nations/

Our commitment to reliable health and medical information. (n.d.). Retrieved on 22 May 2017 from https://www.hon.ch/HONcode/Pro/Visitor/visitor.html

Pauk, W. (1973). The interest level: That's the thing. Journal of Reading, 16(6), 459-461.

Pennbridge, J., Moya, R., & Rodrigues, L. (1999). Questionnaire survey of california consumers' use and rating of sources of health care information including the internet. *Western Journal of Medicine, 171*(5-6), 302-305

Population, total | Data. (n.d.). Retrieved from http://data.worldbank.org/indicator/SP.POP.TOTL

Porter. A, Edirippulige. S; Parents of Deaf Children Seeking Hearing Loss-Related Information on the Internet: the Australian Experience. J Deaf Stud Deaf Educ 2007; 12 (4): 518-529. doi: 10.1093/deafed/enm009

Postel, J. (1994). Domain Name System Structure and Delegation. doi:10.17487/rfc159

Rock, E. L. (1969). A Readability Graph For Russian (doctoral dissertation)

Search engine market share Worldwide | StatCounter Global Stats. (n.d.). Retrieved on 22 May 2017 from http://gs.statcounter.com/search-engine-market-share/all/

Shargorodsky, J., Curhan, S. G., Curhan, G. C., & Eavey, R. (2010). Change in prevalence of hearing loss in US adolescents. Jama, 304(7), 772-778. doi:10.1001/jama.2010.1124

Shepherd, R. K., & Hardie, N. A. (2001). Deafness-induced changes in the auditory pathway: Implications for cochlear implants. Audiology & Neuro-Otology, 6(6), 305-318. doi:10.1159/000046843

Simons, Gary F. and Charles D. Fennig (eds.). 2017. Ethnologue: Languages of the World, Twentieth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Slawinski, E. B., Hartel, D. M., & Kline, D. W. (1993). Self-reported hearing problems in daily life throughout adulthood. Psychology and Aging, 8(4), 552-561. doi:10.1037/0882-7974.8.4.552

Smith, A. (1998). The world health organisation and the prevention of deafness and hearing impairment caused by noise. Noise & Health, 1(1), 6.

Smith, A. (2003). Preventing deafness—an achievable challenge. the WHO perspective. International Congress Series, 1240, 183-191. doi:10.1016/S0531-5131(03)00960-9

Statistics ITU. (n.d.). Retrieved on 27 March 2017. Retrieved from

http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx

Stephens, C. (2000). All about readability.

Stevens, G., Flaxman, S., Brunskill, E., Mascarenhas, M., Mathers, C. D., & Finucane, M. (2013). Global and regional hearing impairment prevalence: An analysis of 42 studies in 29 countries. European Journal of Public Health, 23(1), 146-152. doi:10.1093/eurpub/ckr176

Strawbridge, W. J., Wallhagen, M. I., Shema, S. J., & Kaplan, G. A. (2000). Negative consequences of hearing impairment in old age: A longitudinal analysis. The Gerontologist, 40(3), 320-326. doi:10.1093/geront/40.3.320

Summerfield, Q. (1992). Lipreading and audio-visual speech perception. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 335(1273), 71-78. doi:10.1098/rstb.1992.0009

Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. Journalism Quarterly, 30, 415–433.

Taylor WL. Recent developments in the use of “cloze procedure.” Journal Q. 1958;42-8.

Thorndike, E. L. 1921. The teacher’s word book. New York: Bureau of Publications, Teachers College, Columbia University.

Ventry, I. M., & Weinstein, B. E. (1982). The hearing handicap inventory for the elderly: A new tool. Ear and Hearing, 3(3), 128-134. doi:10.1097/00003446-198205000-00006

Walsh, T. M., & Volsko, T. A. (2008). Readability assessment of internet-based consumer health information. Respiratory Care, 53(10), 1310.

Wang, L., Miller, M. J., Schmitt, M. R., & Wen, F. K. (2013). Assessing readability formula differences with written health information materials: Application, results, and recommendations. Research in Social and Administrative Pharmacy, 9(5), 503-516. doi:10.1016/j.sapharm.2012.05.009

Website | Definition of Website by Merriam-Webster. (n.d.). Retrieved from https://www.merriam-webster.com/dictionary/website

Weinstein, B. E. (1994). Age-related hearing loss: How to screen for it, and when to intervene. Geriatrics, 49, 40–45.

World Internet Users Statistics and 2017 World Population Stats. (n.d.). Retrieved from http://www.internetworldstats.com/stats.htm

Yueh, B., Shapiro, N., MacLean, C. H., & Shekelle, P. G. (2003). Screening and Management of Adult Hearing Loss in Primary Care. JAMA, 289(15), 1976. doi:10.1001/jama.289.15.1976

Appendices

Appendix A

Question posed to informants in the Russian language:

> “Мне нужна помощь с диссертацией, и я надеялся, что вы сможете участвовать. Я смотрю, какие слова люди используют в поиске Google, чтобы искать информацию о чем-либо связанном с слухом в интернете. если бы я спросила: Если вы хотите найти информацию связанно с слухом в интернете, что бы вы искали в Google?”

Appendix B- how to calculate the readability formulas

SMOG calculation (McLaughlin, 1969)-

SMOG RGL = 3 + square root of polysyllable count

F-K calculation (Dubay, 2004)-

Formula:
GL = (.39 x ASL) + (11.8 x ASW) – 15.59

Simplified:
GL = (.4 ASL) + (12 ASW) – 15

Where:
GL = grade level
ASL = average sentence length (the number of words divided by the number of sentences).
ASW = average number of syllables per word (the total number syllables in the sample divided by the number of words).

Dale-Chall calculation (Dubay, 2004)-

1. Select 100-word samples throughout the text (for books, every tenth page is recommended).

2. Compute the average sentence length in words.

3. Compute the percentage of words outside the Dale list of 3,000 words.

4. Compute the following equation:
Score = .1579PDW + .0496ASL + 3.6365

Where: Score = reading grade of a reader who can answer one-half of the test questions on a passage.
PDW= Percentage of Difficult Words (words not on the Dale-Chall word list)
ASL = Average Sentence Length in words.

Dale and Chall also published the following chart for correcting the grade-level scores at the higher grades.

Formula score	Corrected grade level

4.9 and below	Grade 4 and below
5.0 to 5.9	Grades 5-6
6.0 to 6.9	Grades 7-8
7.0 to 7.9	Grades 9-10
8.0 to 8.9	Grades 11-12
9.0 to 9.9	Grades 13-15 (college)
10 and above	Grades 16 and above (college graduate)

Coleman-Liau calculation (Coleman-Liau, 1975)-

Estimated cloze %
= 141.8401 - .214590 L + 1.079812 S

Where: estimated cloze % = percentage of deletions that can be filled in by a college undergraduate
L = number of letters per 100 words
S = number of sentences per 100 words

Grade level
= 27.4004 estimated cloze % + 23.06395

Appendix C- ethics approval letter

HUMAN ETHICS COMMITTEE

Secretary, Rebecca Robinson

Telephone: +64 03 369 4588, Extn 94588

Email: human-ethics@canterbury.ac.nz

Ref: HEC 2017/41/LR

19 June 2017

Alina Filatova
Communication Disorders
UNIVERSITY OF
CANTERBURY

Dear Alina

Thank you for submitting your low risk application to the Human Ethics Committee for the research proposal titled "Evaluation of Online Hearing-related Information in Russian".

I am pleased to advise that this application has been reviewed and approved; **subject to the following**:

please include your name and email address in the top left hand corner of the information sheet.

With best wishes for

your project. Yours

sincerely

pp. R. Robinson

Associate Professor Jane Maidment

Chair, Human Ethics Committee

University of Canterbury Private Bag 4800, Christchurch 8140, New Zealand. **www.canterbury.ac.nz**

Appendix D- Readability scores of the 39 webpages

URL	F-K	Dale-Chall	Coleman-Liau	SMOG	Mean RGL
https://www.phonak.com/ru/ru/poterya-sluha.html	9.11	7.41	8.99	8.48	8.735
http://narodreceptzdorov.ru/bol-v-ushax/uhudshenie-sluha-lechenie	10.39	8.3	10.14	9.24	9.69
http://www.menslife.com/health/26151-poterya-slukha-v-molodosti-prichiny-i-resheniya.html	10.5	9.19	10.51	10.51	10.505
1http://www.oticon.com.ru/hearing/facts/hearing-loss/types-of-hearing-loss.aspx	11.22	9.82	10.69	10.57	10.63
1http://www.news-medical.net/health/Treatment-of-hearing-loss-(Russian).aspx	12.6	9.02	11.76	10.29	11.025
1http://beautyinfo.com.ua/m0c3i3659.html	12.95	9.32	12.15	10.08	11.115
1http://xn--k1agg.net/?view=hearing-16	12.16	10.23	11.78	11.39	11.585
http://narmedzona.ru/zabolevanija-ushej/78-snijenije-sluha	14.08	10.73	13.65	10.99	12.32
https://www.bankreceptov.ru/medsovet/medsovet-0073.shtml	13.8	11.03	12.87	11.78	12.325
https://www.widex.ru/ru-ru/hearing-loss/types-of-hearing-loss/hearing-loss-in-one-ear	13.01	11.88	11.41	13.42	12.445
http://zdravo.by/article/5808/prichiny-ukhudsheniya-slukha	13.6	11.56	13.3	12.21	12.755
1http://www.ayzdorov.ru/lechenie_tygoyxostj_chto.php	14.67	11.12	13.75	12.2	12.975
https://msluh.ru/you-must-tell/vidy-i-stepeni-narushenij-sluxa/	14.69	13.14	12.99	11.96	13.065
http://www.renescenter.ru/poterya_sluha.htm	16.04	11.62	13.97	12.55	13.26
https://www.fdoctor.ru/simptom/snizhenie_slukha/	14.53	12.17	13.7	13.52	13.61
https://naran.ru/bolezni/bolezni-ukha/narusheniya-slukha/	16.4	12.59	14.38	13.58	13.98
https://life5plus.ru/domashnij-doktor/bolezni-organov-dy-haniya-uha-gorla-i-nosa/snizhenie-sluha.html	16.13	12.41	15.76	11.66	14.085
http://xn--k1agg.net/?view=hearing-15	16.8	11.5	15.96	12.55	14.255
http://medportal.ru/enc/otolaryngology/hearloss/	17.59	12.86	16.14	12.74	14.5
http://xn--k1agg.net/?view=hearing-18	17.21	13.51	15.59	13.59	14.59
http://www.dzirde.lv/ru/sluh/vidy-i-stepeni-narushenij-sluha/	17.31	13.96	15.23	11.75	14.595
http://medicalhandbook.ru/disease/3886-poterya-slukha-i-glukhota-tugoukhost.html	17.4	13.19	14.25	15.07	14.66
http://www.doc-plus.ru/uslugi/text/snizhenie-slukha1/	17.67	13.96	15.4	12.7	14.68

http://usharik.ru/site/index.php/sluh-app.html	16.55	14.7	15.2	14.15	14.95
http://www.likar.info/lor/article-66391-vozrastnoe-uhudshenie-sluha/	18.58	14.3	16.59	13.73	15.445
http://www.renescenter.ru/narusheniya_sluha.htm	18.23	14.84	15.75	15.36	15.555
http://slishim.ru/index/problemyi-so-sluxom/lechenie-narushenij-i-problem-so-sluxom/	18	14.62	16.65	14.63	15.64
https://www.stomed.ru/lechenie/lor/poterja-sluha.php	18.37	14.68	16.39	14.91	15.65
http://www.oticon.com.ru/children/facts/hearing-loss/types-of-hearing-loss.aspx	17.49	13.75	16.99	14.79	15.89
http://ilive.com.ua/health/snizhenie-sluha_89597i15995.html	18.19	14.83	16.26	15.53	15.895
http://lookmedbook.ru/disease/snizhenie-sluha	19.13	14.51	17.8	14.29	16.155
Wikipedia	18.67	15.27	17.36	15.2	16.315
http://ilive.com.ua/health/narushenie-sluha_106283i16004.html	19.48	15.48	17.42	14.87	16.45
https://www.bestsound-technology.ru/children/understanding/causes/	20.04	15.04	16.67	16.64	16.655
http://www.gnresound.ru/	20.23	15.18	17.18	16.91	17.045
http://www.who.int/mediacentre/factsheets/fs300/ru/	19.64	16.17	16.72	18	17.36
http://xn--k1agg.net/?view=hearing-loss	20.79	17.01	15.52	19.15	18.08
http://aurora.ua/ru/sluh/diagnostika-sluha_sluh/narushenie-sluha---chto-eto-takoe_.htm	24.66	17.95	21.96	16.87	19.955
http://www.deafworld.ru/poleznaj_info/Vidi_naruheniy_sluha/	23.44	19.01	18.56	21.16	20.085